

anales de psicología, 2016, vol. 32, nº 1 (enero), 256-265
http://dx.doi.org/10.6018/analesps.32.1.194461

© Copyright 2016: Servicio de Publicaciones de la Universidad de Murcia. Murcia (España)
ISSN edición impresa: 0212-9728. ISSN edición web (<http://revistas.um.es/analesps>): 1695-2294

Underlying processes behind false perspective production

Antonio L. Manzanero^{1*}, Beatriz López² & Javier Aróztegui¹

¹ Universidad Complutense de Madrid (Spain), ² University of Portsmouth (U.K.)

Título: Procesos subyacentes más allá de la producción de perspectiva falsa.

Resumen: El objetivo del presente estudio consistió en determinar si los procedimientos de análisis de contenido basados en las propuestas de Reality Monitoring (RM) pueden aportar información útil para discriminar entre declaraciones verdaderas y falsas. Se pidió a los participantes que describieran un accidente de tráfico desde su papel real de testigo o que lo describieran simulando ser la víctima. Los datos se analizaron en función de la exactitud y calidad de los relatos, y se representaron gráficamente mediante Visualización Hiperdimensional (HDV). En el experimento 1 (inter-sujetos) los relatos de las víctimas simuladas contenían más alusiones a procesos cognitivos, más auto-referencias y menos cambios de orden. En el experimento 2 (intra-sujetos) los relatos de las víctimas simuladas también contenían más alusiones a procesos cognitivos y más auto-referencias, pero además fueron menos exactos y aportaron menos información irrelevante y más juicios y comentarios personales. Los gráficos HDV indican que los relatos falsos difieren holísticamente de los reales.

Palabras clave: Memoria autobiográfica; detección de mentira; psicología del testimonio; credibilidad; análisis de contenido; visualización hiperdimensional; escalado multidimensional; holístico.

Abstract: This study aimed to determine the extent to which Reality Monitoring (RM) content analysis can provide useful information when discriminating between actual versus false statements. Participants were instructed to either describe a traffic accident as eyewitness actual role or to describe the accident as a simulated victim. Data were analysed in terms of accuracy and quality, and were represented using high dimensional visualization (HDV). In Experiment 1 (between-participant design), participants made significantly more references to cognitive operations, more self-references and less changes in order when describing the event as simulated victim. In Experiment 2 (within-participants design) participants also made significantly more references to cognitive operations and more self references when describing the event from the simulated victim as well as being less accurate, providing less irrelevant information and more evaluative comments. HDV graphics indicated that false statements differ holistically from actual ones.

Key words: Autobiographical memory; deception detection; eyewitness testimony; credibility; criteria content analysis; high dimensional visualization; multidimensional scaling; holistic.

Introduction

In order to distinguish the origin of memories, people would undertake a process of comparing the characteristic attributes of memories of different origins. This process has been called *Reality Monitoring* (Johnson & Raye, 1981), or in a more generally forms *Source Monitoring* (Johnson, Hashtroudi & Lindsay, 1993). Since the proposal of the Reality Monitoring (RM) model and the use of phenomenological content analysis to differentiate between memories from different sources, there have been numerous studies investigating distinctive characteristics of memories from a wide variety of sources such as actual memories from perceptual processes, and self-generated memories from imagination, dreams, fantasies, false memories and post-event implanted memories. In their original proposal, Johnson and Raye (1981) suggested that perceptual memories would have more contextual, sensorial and semantic attributes whilst self-generated memories would have more information about cognitive operations. Subsequent research (Manzanero & Diges, 1995; Manzanero, 2006; Santtila, Roppola, & Niemi, 1998; Schooler, Gerhard, & Loftus, 1986; Strömwall, Bengtsson, Leander, & Granhag, 2004; Vrij, 2005) has identified additional attributes to include, amongst others, doubts, irrelevant information, explanations, self-references, exaggerations, personal involvement, judgements and comments, pauses, spontaneous corrections, changes in order, length of description... whilst other researchers have proposed alterna-

tive criteria to evaluate the credibility of testimonies in specific criminal cases (Jones & McGraw, 1987; Steller & Köhnken, 1989; Trankell, 1972; Undeutsch, 1989).

However, the evidence about potential differences between true and false statements is largely inconsistent. Vrij, Akehurst, Soukara and Bull (2004) found significant differences in the expected direction regarding logical structure, amount of details, contextual information, descriptions of interaction, conversation reproduction, visual, audio, spatial and temporal details and cognitive operations. No effects were found however regarding mental state attributions, acknowledgement of lack of memory and doubts. Other studies however fail to find such marked differences. Porter and Yuille (1996) analysed differences between the two types of statements using some of the content criteria from Criteria-Based Content Analysis (CBCA) procedure for credibility assessment (Köhnken, Manzanero, & Scott, 2015; Steller & Köhnken, 1989). Out of the 17 criteria employed only three yielded significant differences: amount of detail provided, logical structure and acknowledgement of lack of memory. Although the CBCA came independently to the studies on the RM processes, assumptions are shared, and both have been proposed as a method to assess the credibility of the statements with similar results, albeit RM could be more sensitive than CBCA (Granhag, Strömwall, & Landström, 2006). Sporer and Sharman (2006) analysed made-up or true autobiographical data according to clarity, how vivid they were, sensory, spatial and temporal information, quality of the memory, significance and realism, in the theoretical framework of RM. Results showed significant differences in the predicted direction only for temporal information and realism.

* Dirección para correspondencia [Correspondence address]:

Antonio L. Manzanero. Facultad de Psicología. Universidad Complutense de Madrid. Campus de Somosaguas. CP 28223. Madrid (España).
E-mail: antonio.manzanero@psi.ucm.es

Different reasons can be argued to explain the inconsistency of the results and the lack of differences between real and false statements. a) Mounting evidence (Bensi, Gambetti, Nori, & Giusberti, 2009; Campos & Alonso-Quecuty, 1998; Colwell, Hiscock-Anisman, Memon, Colwell, Taylor, & Woods, 2009; Comblain, D'Argembeau & Van der Linden, 2005; Diges, 1995; Diges, Rubio & Rodríguez, 1992; Henkel, Franklin & Johnson, 2000; Manzanero, El-Astal & Aróztegui, 2009; Memon, Fraser, Colwell, Odinet & Mastroberardino, 2010; Pérez-Mata & Diges, 2007; Roberts & Lamb, 2010) suggests that numerous factors (activation levels, coherence and previous knowledge, perceptual modality, subjects' involvement degree and perspective, age, contextual factors, type of design used in the research, training of the evaluators, and the coding method) influence the characteristics of statements. b) There is also great variability in the possible origins of memories and different sources produce distinctive characteristics which vary depending on whether they are from fantasy, lies, dreams or post-event information (Hekkanen & McEvoy, 2005; Johnson, Kahan & Raye, 1984; Sporer & Sharman, 2006; Vrij et al., 2004). c) Moreover, within each source there are different degrees on a continuum from perceptual memories to the most fantastic memories. For instance it is not the same to imagine a complete event than to change a detail of a perceptual memory, for example when a witness pretend to have been the victim, or a perpetrator pretend to be a witness. In these cases, the person has experienced the event and has all the information pertaining to this event but his role is very different from the one he is portraying. When adopting a false role, the accuracy and phenomenological characteristics of the description of an event may be affected. According to the RM model (Johnson & Raye, 1981; Johnson et al., 1993) the statements from real witness should be characterised by attributes associated to perceptual memories whilst witnesses adopting a false role should provide a very different statement characteristic of self-generated memories. In a seminal study Anderson and Pichert (1978) asked participants to remember a story from the perspective of a prospective buyer or a prospective burglar. Results showed that participants remembered different details from each perspective and so they concluded that having a specific schema may affect encoding as well as recovery and they also suggested that different relevance criteria determine whether details are accessible or not, even when they all are stored in memory.

The present study aimed to further our understanding of the differences between actual and false statements by analysing the attributes of statements produced by witnesses that supplant the role of the victim of an event. More specifically, actual statements given by observers were compared with statements given by simulated victims (eyewitnesses who pretend to be a victim). Therefore, false accounts are provided by witnesses who play a role that is false, and knowing that this is not true they intent to deceive in their testimonies.

According to the RM model it was predicted that actual statements would present more sensorial and contextual information, less references to cognitive processes, less doubts, explanations and self-references and would be shorter than false statements. Other attributes that could be of use to discriminate between the two types of statements such as spontaneous corrections, unstructured production, exaggerations, and judgements and personal comments will also be explored in this study. Two alternative theoretical perspectives will be tested: a general lying theory will be compared to a differential lying theory. The general lying theory states a unique, general, way to build the lying statement, leading to homogeneous differences across participants in statements when compared with those not lying. On the other hand, the differential lying theory states that each person uses its own set of believes when trying to solve the problem on how to build a credible lying statement. Different beliefs may lead to differences with respect to those not trying to lie, but the differences may be heterogeneous. For example, if a person believes that a more detailed statement is more credible, it will likely show more details than those not trying to lie, while another person believing that too many details looks like a prepared statement, may produce less details than those not trying to lie.

When attempting to discriminate true and false memories we are faced with two possibilities. The first scenario is that some of the variables can significantly and consistently discriminate between both types of memories, in which case data analyses of paired comparisons are appropriate. A second alternative however is that the distinction between the two types of memories is only possible when all variables are considered simultaneously in a holistic way. Thus in this study in addition to utilising paired comparisons of accuracy and quality measures, high dimensional visualization (HDV) was employed to represent and explore complex patterns of interactions between variables. This technique is a multivariate analysis used in psychology (Clark-Carter, 2004; Manzanero et al., 2009; Manzanero, Alemany, Recio, Vallet & Aróztegui, 2015; Manzanero, López, Aróztegui, & El-Astal, 2015; Roskam, 1989; Tversky & Krantz, 1970) and other disciplines (Barton & Valdés, 2008; Romero, Valdés & Barton, 2007) to predict outcomes from a large set of variables. If discrimination rates of true and false statements are more successful analysing patterns of interactions than pairwise comparisons it would suggest that it may be time to discard simpler general models of credibility assessments in favour of more complex differential theoretical models of deception.

Experiment 1

Method

Design

An unifactorial design was employed using reality of description as the between-participant factor with two levels: a) actual witness and b) simulated victims. There were two categories of dependent measures: accuracy and qualitative measures. Considering the RM theoretical framework, accuracy measures include *correct information*, amount of correct details provided; and *distortions*, incorrectly described details or fabrications.

Qualitative measures include *sensory information*, information relating to sensory aspects of the event: colours, sizes...; *contextual information*, information about spatial and temporal aspects of the environment in which the event takes place; *cognitive operations*, explicit allusion to cognitive operations; *doubtful expressions*, expressions that indicate doubts about the description in hand (i.e., it could be, it appears, I think, it is probable...); *irrelevant information*, correct information that does not form part of the pre-established script. It is used for the overall correct information variable; *explanations*, information that adds to purely sensorial or contextual information, such as, for instance, a functional description of some aspect of the event; *spontaneous corrections*, corrections made by the participant that appear crossed out, corrected or added to the text; *change in order*, changes in the order in which the event took place; *exaggerations*, descriptions in which the quality of the events is over or under-reported; *judgements and personal comments*, judgements about some aspect of the event and personal additions; *self-references*, number of first person pronouns or first person verb forms that revealed a pronoun in the deeper structure; and *length of the narrative*, the number of words of the report.

Participants

Thirty psychology university students (21 female and 9 male) were randomly selected and allocated to two groups corresponding to the two types of descriptions under scrutiny. The ages ranged between 17 and 27 years old, with mean age of 19.16 years ($SD = 2.61$).

Materials

A complex event employed in previous studies (e.g. Manzanero & Diges, 1995; Manzanero, 2009) was presented to participants. This event was presented as a short film, 29 sec. in length and without sound. This consisted of a cross roads collision between two cars. The clip starts with a car moving amongst other vehicles in a street by a park. The car arrives to a cross roads, stops and immediately starts again and collides at low speed with another vehicle coming from the side. The end of the film shows the outcome of the col-

lision, the first vehicle has been pushed along the cross roads and is seriously damaged.

All participants were told in advance about the nature of the event; of its brevity and lack of sound. They watched the film in a colour TV monitor in a classroom at the university. The recall of the event took place in the same lecture theatre. An analysis protocol was used to measure the accuracy of descriptions provided by the study subjects. This protocol describes the event through propositions, and it has been shown to be useful in accuracy assessment, to avoid bias and to facilitate scoring (Manzanero & Diges, 1995).

Procedure

Participants were asked to watch carefully the TV monitor where a short clip showing a traffic accident was presented. Immediately after watching the clip, participants performed a distractor task for 10 minutes, which involved describing an itinerary on the university campus from the railway station to the School of Psychology. Once completed, participants were asked to describe the clip. Fifteen participants were asked to describe it as much detail as possible as if they were the victim of the accident and fifteen participants were asked to remember as much detail as possible from their actual role as eyewitness.

Data Analysis

Reliability. The protocols of the free narrative account were submitted to two expert coders for content analysis. The expert coders, were trained specially for this study with examples taken from each category. To assess the codings' reliability for within- and between-coders, the Agreement Index [$AI = \text{agreements} / (\text{agreements} + \text{disagreements})$] was computed. In all of the variables measured, this was greater than the cut-off .80 (Iversky, 1977).

HDV graphs: To explore the holistic complex patterns that underlie the construction of lies, HDV graphs were used, considering all content criteria simultaneously (Manzanero et al., 2009). This technique is linked, at its root, to a scatter plot. The different variables corresponding to each statement are represented as a point in a high dimensional space. When having more than three variables, such as it is the case in this study, mathematical techniques are required to reduce dimensionality to allow a 3D graphical representation. This study employed multidimensional scaling in order to reduce dimensionality (Buja, Swayne, Littman, Dean, Hofmann & Chen, 2008; Steyvers, 2002). Each point in the hyperspace has a distance to each other point. Multidimensional scaling will search 3D points preserving those distances between points as much as possible. Distance between points will be calculated using a normalized euclidean distance (Barton & Valdés, 2008) as shown below:

$$d_{ij} = \sqrt{(1/p) \sum_{j=1}^p (x_{ij} - t_{ij})^2}$$

The normalized euclidean distance is equivalent to regular euclidean distance but the apparent higher distance due

$$\text{Sammon's Error} = \frac{1}{\sum_{i < j} \delta_{ij}} \cdot \frac{\sum_{i < j} (\delta_{ij} - \zeta_{ij})^2}{\delta_{ij}}$$

Where δ_{ij} is the distance (or dissimilarity) between points i and j in the original space and ζ_{ij} is the distance in the reduced dimensionality 3D space. Therefore, Sammon's Error compares the differences between the original distances between points, in the original hyperspace, and the final distance among the new representing points in the 3D space. If distances are similar, the error is small as few information is lost in the transformation. If distances are high, the error is high as more information is lost in the transformation. Sammon's error value should be interpreted as the proportion of distance difference with respect to the original distance. For example, .05 means that the average difference between distances is 5 percent with respect to the original distance.

Finally, the 3D points are represented using Virtual Reality Modelling Language (VRML). VRML files allow graph-

to a high number of dimensions is corrected through normalization (that is, considering p , the number of dimensions). The quality of the 3D transformation is measured through the Sammon's Error (Barton & Valdés, 2008; Romero, Valdés & Barton, 2007) as follows:

ical rotation and exploration to facilitate graphical data analysis. Two HDV graphs were used in each of the experiments. The first HDV graph represented only the variables that yielded significant differences in the statistical analyses, whilst the second HDV included all variables regardless of whether they were significant in the pairwise statistical comparisons. Finding that there was a significant increase in classification power with all the variables included would suggest that valuable information may be lost by using the traditional classification system of statistical pairwise comparisons.

Results

Once statements were transcribed, they were analysed according to the categories described above. Results for accuracy and qualitative measures are reported separately.

Accuracy measures

No significant group differences were found in accuracy measures. See Table 1 for a summary of results.

Table 1. Average scores (and standard deviations) in accuracy and qualitative measures in Experiment 1.

Type of account	Actual witness	Simulated victim	Effect Size (r)
Accuracy details	14.66 (3.35)	14.18 (4.03)	.064
Distortions	1.40 (1.59)	1.46 (1.64)	-.018
Sensory information	4.33 (2.09)	3.60 (1.50)	.196
Contextual information	7.33 (3.98)	7.86 (3.81)	-.067
Cognitive operations*	.60 (.98)	2.20 (1.85)	-.475
Hesitant expressions	.60 (.74)	.26 (.46)	.265
Spontaneous corrections	2.66 (2.53)	1.66 (1.76)	.223
Change of order*	.40 (.51)	.06 (.26)	.387
Explanations	1.73 (1.39)	1.13 (1.30)	.217
Exaggerations	.73 (.80)	.53 (.64)	.136
Irrelevant information	1.13 (1.30)	.46 (.52)	.320
Judgements and comments	2.13 (1.77)	1.46 (1.64)	.192
Self-references*	.46 (1.06)	9.20 (4.52)	-.799
Length	153.46 (61.48)	123.60 (49.54)	.258

* Significant differences

Qualitative measures

Significant differences were found on self-references, cognitive operations and change of order. Participants describing the event as simulated victim had more self-references ($t(28)=7.281$, $p<.01$, $r_{\text{v} \lambda}=.809$), more cognitive operations ($t(28)=2.945$, $p<.01$, $r_{\text{v} \lambda}=.486$), and less changes

of order ($t(28)=2.269$, $p<.05$, $r_{\text{v} \lambda}=.394$), than those describing the event as eyewitness. Results can be seen in Table 1.

The HDV results are shown in Figures 1a and 1b. Figure 1a represents the three variables that yielded significance in the statistical analyses. The classification obtained using only these three variables was not optimal, with only 80 percent of cases correctly classified.

Figure 1b represents all variables regardless of whether their values in both groups were significantly different in the

statistical analyses. The quality of the dimensionality reduction through MDS was very good, high with a small sammon's error (sammon's error = .027). In this case, the dotted line graphically discriminating actual witness and simulated victims, provides a correct classification in 96.7 percent

of cases, a remarkably better discrimination than the classification obtained just using the statistically significant variables.

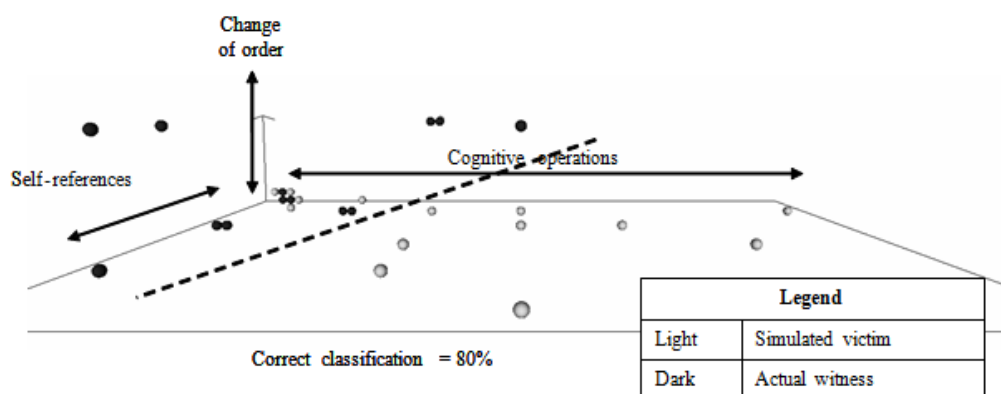


Figure 1a. HDV graph including statistically significant variables only in Experiment 1. No sammon error is shown as no dimensionality reduction was needed with only 3 variables. The dotted line indicates the best graphical classification of data.

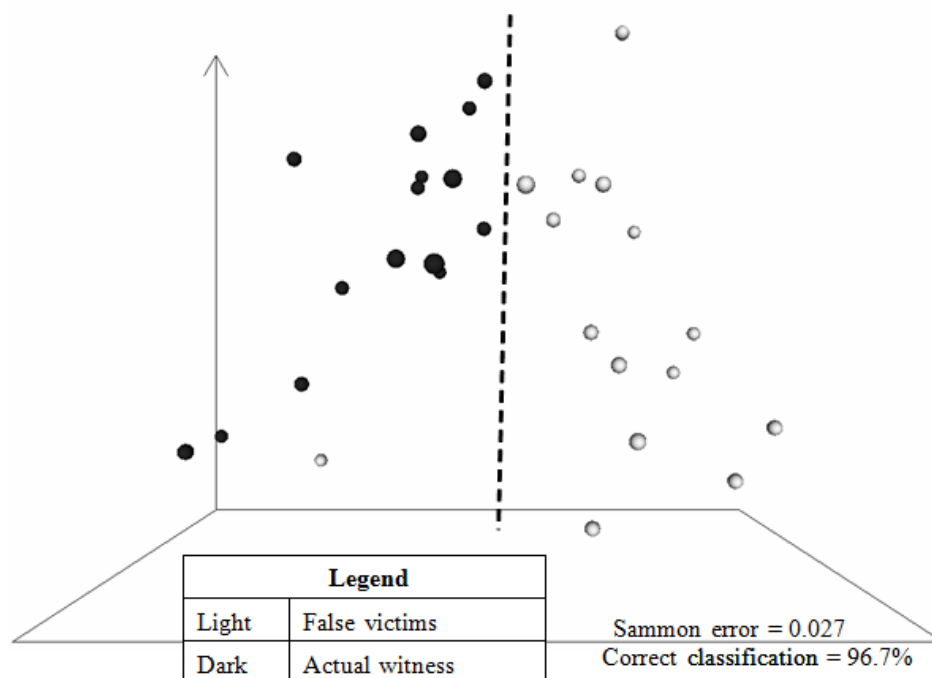


Figure 1b. HDV graph including all variables in used in Experiment 1 (between-participants). Length of narrative was normalized with respect to the rest of variables in order to prevent a distortion in data interpretation due to just one variable.

Discussion

Only a few significant differences were found between actual and false statements. As predicted, participants made significantly more references to cognitive operations and more self-references when describing the event as simulated

victim. Contrary to predictions simulated victims made fewer changes in order in the perceptual condition.

One possible explanation for the lack of differences in the statistical analyses may relate to the use of a between-participants design. In a previous study, Bensi and cols. (2009) have shown that the methodological design, whether within or between-participants, influences credibility dis-

crimination. Specifically, they found significant differences between truth and deception using within-participants design but not with a between-participants design. Undeutsch (1989) suggests that to analyse credibility of statements in terms of content analyses it is necessary to compare a statement known to be true with the target statement by the same individual. A second experiment was therefore designed to compare descriptions of the same participants under the two conditions of actual witness and false victims to enable us to answer this question. It was expected that this manipulation would increase the differences between conditions on both accuracy and quality measures.

Experiment 2

Method

Design

The same design as in Experiment 1 was employed in this experiment, except that the independent variable (whether the statement was actual or false) was tested with a within-participant design. The same dependent variables, accuracy and quality, were measured.

The same materials and procedure as in experiment 1 were used.

Participants

Thirty-five psychology university students were randomly selected and assigned to two groups. The first group comprised of 18 participants and they described the clip as simulated victim first followed by the description as actual witness. The remaining 17 participants described the event in the reverse order, first as actual witness and then as simulated victim.

Data Analysis

The same reliability procedures and HDV representations were used as in Experiment 1. Equally, to assess the codings' reliability for within- and between-coders, the Agreement Index was computed. In all of the variables measured, this was greater than the cut-off .80 (Tversky, 1977).

Results

As in the previous experiment, once descriptions were transcribed they were analysed according to the categories mentioned earlier. Results are also presented separately for accuracy and quality measures.

Accuracy measures

Statements were significantly more accurate in the observer condition than in the simulated condition ($t(34)=3.479$, $p<.01$, $r_{\lambda}=.269$). No differences were found for distortions between the two conditions. See Table 2 for a summary of results.

Table 2. Average scores (and standard deviations) in accuracy and qualitative measures in Experiment 2.

Type of account	Actual witness	Simulated victim	Effect Size (r)
Accuracy details*	14.94 (3.51)	12.71 (4.42)	.269
Distortions	1.48 (1.98)	1.71 (2.44)	-.051
Sensory information	3.91 (2.34)	3.28 (3.93)	.097
Contextual information	7.94 (4.71)	7.28 (4.32)	.073
Cognitive operations*	.82 (1.65)	3.08 (2.56)	-.465
Hesitant expressions	.42 (.65)	.51 (1.01)	-.053
Spontaneous corrections	1.88 (2.03)	2.08 (2.76)	-.041
Change of order	.20 (.41)	.05 (.24)	.217
Explanations	1.20 (1.28)	1.23 (1.11)	-.012
Exaggerations	.57 (.70)	.57 (.70)	0
Irrelevant information *	1.02 (1.27)	.40 (.65)	.294
Judgements and comments*	1.34 (1.61)	2.77 (2.74)	-.303
Self-references*	.60 (1.06)	11.68 (3.61)	-.901
Length	139.00 (66.59)	147.68 (62.04)	-.067

*Significant differences

Qualitative measures

Significantly more self-references, less irrelevant information and more judgements and personal comments, and references to mental processes were found in false statements than in actual ones ($t(34)=11.423$, $p<.01$, $r_{\lambda}=.901$; $t(34)=3.061$, $p<.01$, $r_{\lambda}=.294$; $t(34)=3.515$, $p<.01$, $r_{\lambda}=.303$; $t(34)=4.218$, $p<.01$, $r_{\lambda}=.465$, respectively). For a summary of results see Table 2.

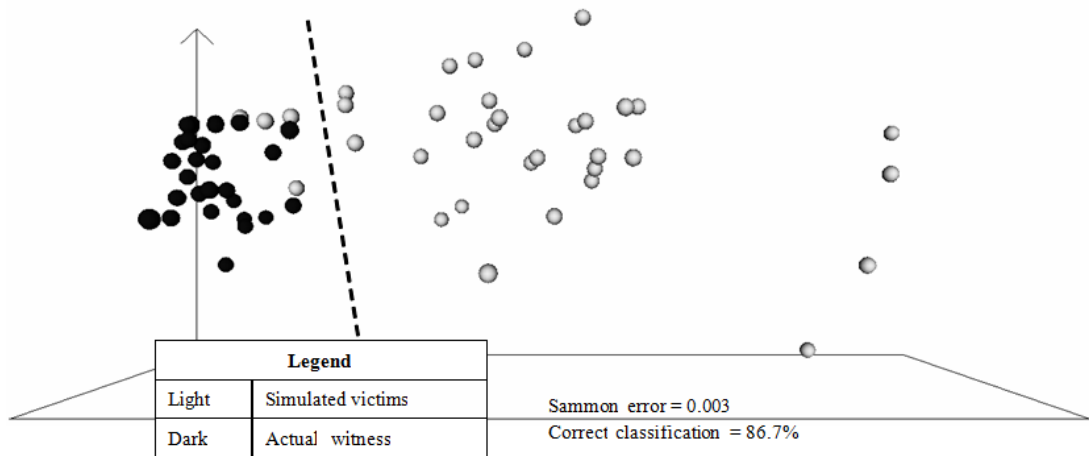


Figure 2a Within-participants HDV graph including statistically significant variables only.

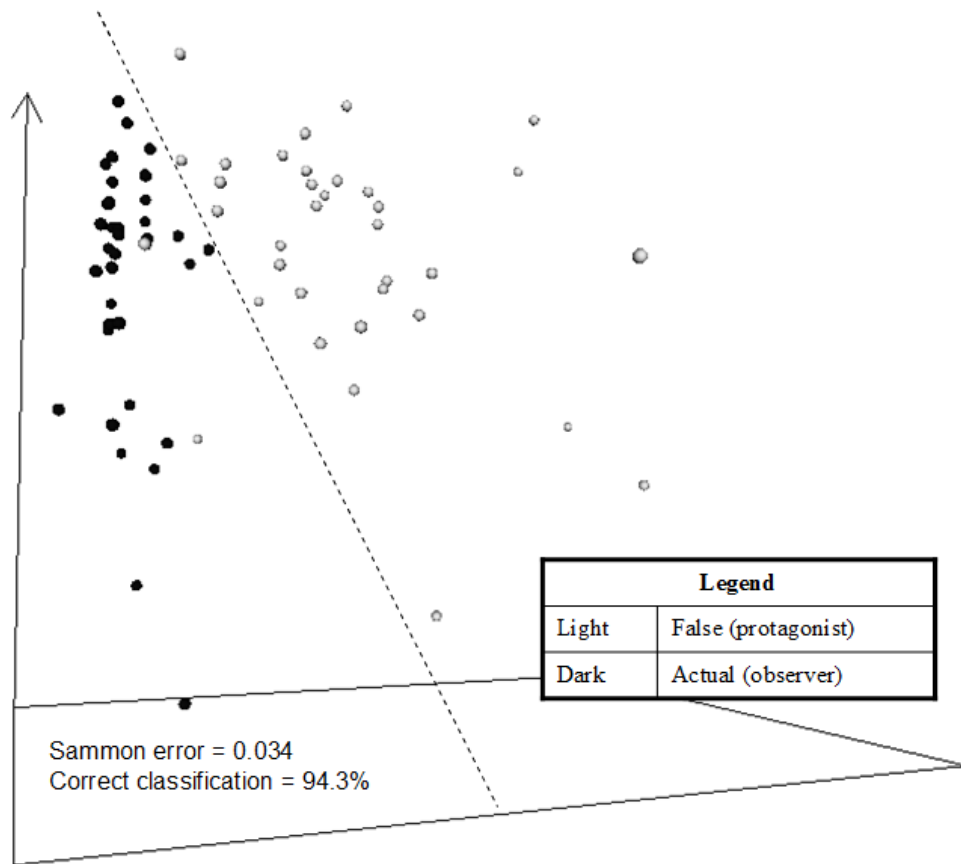


Figure 2b. HDV graph including all variables in used in Experiment 2 (within-participants). Length of narrative was normalized with respect to the rest of variables in order to prevent a distortion in data interpretation due to just one variable.

As in the previous Experiment, a HDV representation was employed using, first, only the four variables that were significant in the statistical analyses (see Figure 2a). The quality of the dimensionality reduction was very good, with a sammon's error = .003. The plane used to classify (separat-

ing points in 2 groups) permits a value of correct classification of 86.7 percent.

Figure 2b shows the HDV graph including all variables measured in Experiment 2. The dimensionality reduction quality again was very high, with a sammon's error = .034.

When discriminating data using the best possible plane, a 94.3 percent correct classification of statements (true or false) was achieved. This is a significant improvement from the classification achieved by including only those variables that produced significant differences in the statistical analyses.

Discussion

Whilst in Experiment 1, using a between-participants design, no differences were found in accuracy and amount of irrelevant information, using a within-participant design has produced significant differences in both these measures. Simulated victims' statements contained less accurate and less irrelevant information than actual witnesses' statements. In contrast, both studies have showed consistent differences in references to cognitive operations and self-references. Moreover, the within-participant design also produced differences in personal commentaries whilst the between-participant design has produced differences in changes of order. Significant differences found in relation to cognitive operations, are probably due to processes relating to the reconstruction of the event to integrate information with the simulated victim. This reconstruction, according to the RM model (Johnson & Rye, 1981) may be the result of false accounts undergoing a process of internalisation.

Leaving aside the effects due to processes of internalisation, differences between witnesses' and simulated victim's statements, when taking into account the results from classical statistical analysis, are minimal regardless of whether a between- or within-participants design was employed. The failure to discriminate between both statements using pairwise comparisons, seems to be related to the underlying theoretical perspective as in contrast in both experiments HDV representations showed different complex patterns for two types of statements.

Conclusions

It seems rather difficult to discriminate between statements when an eyewitness pretends to be another person involved in the event with a different role, in this case a simulated victim. This adds to existing evidence showing little differences in memories from different sources. The few differences found in adults when recalling emotional events, such as a traffic accident, point to a good ability to generate different credible versions of the same event. The few differences found can also explain as well why it is difficult to judge the credibility of statements, as many other studies show (Santtila et al., 1998; Schooler et al., 1986; Strömwall et al., 2004; Vrij, 2005).

If we apply the content criteria defined by Steller and Köhnken (1989) using the SVA/CBCA technique, it would still be difficult to discriminate between actual and false statements as only three variables showed significant differ-

ences, namely, unstructured production (changes of order in Exp. 1), amount of details provided (accuracy details in Exp. 2) and superfluous details (irrelevant information in Exp. 2). And even though not all criteria can be applied to the recall of a traffic accident, no differences were found on measures relating to contextual embedding, spontaneous corrections, doubts about one's own memory, and admitting lack of memory.

Adding all variables rather than including only statistically significant variables increased the differences in both Experiments 1 and 2. This suggests that a complex pattern, not detectable with statistical pairwise data analysis, seems to underlie simulated victims' statements. Including all variables available, improved classification of actual and false statements from 80 percent to 97.6 percent in Experiment 1 and from 86.7 percent to 94.3 percent in Experiment 2. Similar results were found in previous studies when actual and simulated victims in people with intellectual disability were compared (Manzanero, Alemany et al., 2015).

The conclusions stemming from these findings do not just relate to the use of pairwise comparisons or HDV techniques. The crucial issue relates to the theoretical assumptions behind these differences. If there is a general process behind the production of false statements, and if the process affects the variables measured in a unique way, single criterion should be enough to discriminate between true and false statements.

In contrast, if what underlies the production of false statements is the result of a complex process involving many factors, alternative processes, and individual differences on what makes a false statement look credible, then using pairwise comparisons will not be successful in discriminating between true and false statements. If this is the case, multidimensional scaling techniques need to be applied.

Lying is a general cognitive skill appearing quite early in cognitive development as part of the development of the theory of mind (Feldman, 2006; Flavell et al., 1995; Rochat, 1999; Zahn-Waxler, Robinson, & Emde, 1992; Lee & Homer, 1999). It has been assumed that once children acquire the concept of lying, they start producing lies in a similar way. In contrast, we propose an alternative model, the *differential lying theory*, which suggests that although the comprehension of what is a lie is universal across individuals, throughout development beliefs on what makes a statement look credible diverge among individuals.

Past attempts have been made to objectify as much as possible the evaluation of credibility in adults through the analyses of behavioural and physiological factors related to witness' statements and also through the analyses of the content of these statements. None of these methods however allow the *objective* evaluation of credibility, and less so, that of the veracity, as illustrated by findings of near to 50% of errors in discriminating source of accounts (Manzanero & Diges, 1994; Santtila et al., 1998; Schooler et al., 1986; Strömwall et al., 2004). The findings from the present study suggest that previous attempts may have been constrained

by the comparison of isolated variables rather than investigating more complex patterns which take into account all variables simultaneously. When all variables are taken into account the degree of distinction improved considerably. Perhaps it is the reason because Hartwig & Bond (2011), as a result of a meta-analysis, argued that intuitive notions about deception are more accurate than explicit knowledge and that lie detection is more readily improved by increasing behavioural differences between liars and truth tellers than by informing lie-catchers of valid cues to deception. Intuition judgements can take into consideration more criteria to distinguish between both statements than explicit knowledge

based on content criteria techniques. Anyway, intuition seems to be few effective when accuracy credibility assessment is the goal (Manzanero, Quintana, & Contreras, 2015).

The results from this study however should be taken with caution as it is only a first exploratory study which assesses only one variable -simulated victim- and thus would need further confirmation in future research.

Acknowledgements.- This work is part of a research project entitled *The witness evidence in specially vulnerable victims*, developed by the UCM Research Group on Psychology of Testimony (www.psicologiadeltestimonio.com) (ref. 971672).

References

- Anderson, R. C. & Pichert, J. W. (1978). Recall of previously unrecalled information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, 16, 1-12. doi: 10.1016/S0022-5371(78)90485-1
- Barton, A. & Valdés, J. J. (2008). Hybrid unsupervised/supervised virtual reality spaces for visualizing gastric and liver cancer databases: An evolutionary computation approach. In A. An, S. Matwin, Z. W. Ras & D. Ślezak (Eds.), *Foundations of Intelligent Systems* (pp. 256-261). Berlin: Springer.
- Bensi, L., Gambetti, E., Nori, R. & Giusberti, F. (2009). Discerning truth from deception: the sincere witness profile. *The European Journal of Psychology Applied to Legal Context*, 1(1), 101-121. Retrieved from http://www.usc.es/sepjf/images/documentos/volume_1/Bensi.pdf
- Buja A., Swayne D. F., Littman M., Dean N., Hofmann H. & Chen L. (2008). Data Visualization with Multidimensional Scaling. *Journal of Computational and Graphical Statistics*, 17(2), 444-472. doi:10.1198/106186008X318440
- Campos, L. & Alonso-Quecuty, M. L. (1998). Knowledge of crime context: Improving the understanding of why the cognitive interview works. *Memory*, 6, 103-112. doi: 10.1080/741941602
- Clark-Carter, D. (2004) *Quantitative psychological research*. New York: Psychology Press.
- Colwell, K., Hiscock-Anisman, C., Memon, A., Colwell, L. H., Taylor, L., & Woods, D. (2009). Training in Assessment Criteria Indicative of Deception to improve credibility judgments. *Journal of Forensic Psychology Practice*, 9(3), 199-207. doi:10.1080/15228930902810078
- Comblain, C., D'Argembeau, A. & Van der Linden, M. (2005). Phenomenal characteristics of autobiographical memories for emotional and neutral events in older and younger adults. *Experimental Aging Research*, 31(2), 173-89. doi: 10.1080/03610730590915010
- Diges, M. (1995). Previous knowledge and delay in the recall of a filmed event. In G. Davies, S.M.A. Lloyd-Bostock, M. McMullan & C. Wilson (Eds.), *Psychology, law and criminal justice. International developments in research and practice* (pp. 46-55). Berlin: W. de Gruyter.
- Diges, M., Rubio, M. E. & Rodríguez, M. C. (1992). Eyewitness testimony and time of day. In F. Lösel, D. Bender & T. Bliesener (Eds.), *Psychology and Law. International Perspectives* (pp. 317-320). Berlin: W. de Gruyter.
- Feldman, R. S. (2006). *Development across the life span*. Englewood Cliffs, NJ: Prentice-Hall.
- Flavell, J. H., Green, F. L. & Flavell, E. R. (1995). The development of children's knowledge about attentional focus. *Developmental psychology*, 31, 706-712. doi: 10.1037/0012-1649.31.4.706
- Granhag, P. A., Strömwall, L. A., & Landström, S. (2006). Children recalling an event repeatedly: Effects on RM and CBCA scores. *Legal and Criminological Psychology*, 11(1), 81-98. doi: 10.1348/135532505X49620
- Hartwig, M., & Bond Jr., Ch. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin*, 137(4), 643-659. doi: 10.1037/a0023589
- Hekkanen, S. T. & McEvoy, C. (2005). Source monitoring in eyewitness memory: Implicit associations, suggestions, and episodic traces. *Memory and Cognition*, 33(5), 759-769. doi: 10.3758/BF03193072
- Henkel, L. A., Franklin, N. & Johnson, M. K. (2000). Cross-modal source monitoring confusions between perceived and imagined events. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26(2), 321-335. doi: 10.1037/0278-7393.26.2.321
- Johnson, M. K., Hashtroudi, S. & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3-28. doi: 10.1037/0033-2909.114.1.3
- Johnson, M. K., Kahan, T. L. & Raye, C. L. (1984). Dreams and reality monitoring. *Journal of Experimental Psychology: General*, 113(3), 329-344. doi: 10.1037/0096-3445.113.3.329
- Johnson, M. K. & Raye, C. (1981). Reality monitoring. *Psychological Review*, 88, 67-85. doi: 10.1037/0033-295X.88.1.67
- Jones, D. P. H. & McGraw, J. M. (1987). Reliable and fictitious accounts of sexual abuse in children. *Journal of Interpersonal Violence*, 2, 27-45. doi: 10.1177/088626087002001002
- Köhnken, G., Manzanero, A. L., & Scott, M. T. (2015). Análisis de la Validez de las Declaraciones (SVA): mitos y limitaciones (Statement Validity Assessment: Myths and limitations). *Anuario de Psicología Jurídica*, 25, 13-19. doi:10.1016/j.apj. 2015.01.004
- Lee, K. & Homer, B. (1999) Children as folk psychologists: The developing understanding of the mind. In A. Slater & D. Muir (Eds.), *The Blackwell reader in developmental psychology*. Malden, MA: Blackwell.
- Manzanero, A. L. (2006). Do perceptual and suggested accounts actually differ? *Psychology in Spain*, 10(1), 52-65. Retrieved from http://eprints.ucm.es/11380/1/psychology_in_spain.pdf
- Manzanero, A. L. (2009). Análisis de contenido de memorias autobiográficas falsas (Criteria content analysis of false autobiographical memories). *Anuario de Psicología Jurídica*, 19, 61-72.
- Manzanero, A. L., Alemany, A., Recio, M., Vallet, R., & Aróztegui, J. (2015). Credibility assessment of statements given by people with learning disability. *Anales de Psicología*, 31 (1), 1-25. doi:10.6018/analesps.31.1.166571
- Manzanero, A. L. & Diges, M. (1994). Análisis de la credibilidad de recuerdos percibidos e imaginados (Credibility assessment of perceived and imagined memories). *Apuntes de Psicología*, 41-42, 81-92. Retrieved from: <http://eprints.ucm.es/11043/>
- Manzanero, A. L. & Diges, M. (1995). Effects of preparation on internal and external memories. In G. Davies, S. M. A. Lloyd-Bostock, M. McMullan & C. Wilson (Eds.), *Psychology, law and criminal justice. International developments in research and practice* (pp. 56-63). Berlin: W. De Gruyter & Co. Retrieved from http://eprints.ucm.es/6174/1/oxford_95.pdf
- Manzanero, A. L., El-Astal, S. & Aróztegui, J. (2009). Implication degree and delay on recall of events: An experimental and HDV study. *The European Journal of Psychology Applied to Legal Context*, 1(2), 101-116. Retrieved from http://eprints.ucm.es/11383/1/Manzanero_et_al_2009.pdf
- Manzanero, A. L., López, B., Aróztegui, J., & El-Astal, S. (2015). Autobiographical memories for negative and positive events in war contexts. *Anuario de Psicología Jurídica*, 25, 57-64. doi:10.1016/j.apj. 2015.02.001
- Manzanero, A. L., Quintana, J. M., & Contreras, M. J. (2015). (The null) Importance of police experience on intuitive credibility of people with intellectual disabilities. *Research in Developmental Disabilities*, 36, 191-197. doi:10.1016/j.ridd.2014.10.009

- Memon, A., Fraser, J., Colwell, K., Odinet, G., & Mastroberardino, S. (2010). Distinguishing truthful from invented accounts using reality monitoring criteria. *Legal and Criminological Psychology*, 15(2), 177–194. doi: 10.1348/135532508X401382
- Pérez-Mata, N. & Diges, M. (2007). False recollections and the congruence of suggested information. *Memory*, 15(7), 701–717. doi:10.1080/09658210701647258
- Porter, S. & Yuille, J. C. (1996). The language of deceit: an investigation of the verbal cues to deception in the interrogation context. *Law and Human Behavior*, 20, 443–458. doi: 10.1007/BF01498980
- Roberts, K. P., & Lamb, M. E. (2010). Reality-monitoring characteristics in confirmed and doubtful allegations of child sexual abuse. *Applied Cognitive Psychology*, 24(8), 1049–1079. doi: 10.1002/acp.1600
- Rochat, P. (Ed.) (1999) *Early social cognition: Understanding others in the first months of life*. Mahwah, NJ: Erlbaum.
- Romero, E., Valdés, J. J. & Barton, A. J. (2007). Neural network based virtual reality spaces for visual data mining of cancer data: An unsupervised perspective. In F. Sandoval, A. Prieto, J. Cabestany & M. Graña (Eds.), *Computational and ambient intelligence*. (pp. 1020–1027). Berlin: Springer.
- Roskam, E. E. (1989) Formal models and axiomatic measurement. In E. E. Roskam (Ed.), *Mathematical psychology in progress*. Berlin: Springer-Berlag.
- Santtila, P., Roppola, H. & Niemi, P. (1998). Assessing the truthfulness of witness statements made by children (aged 7–8, 10–11, and 13–14) employing scales derived from Johnson and Raye's model of reality monitoring. *Expert Evidence*, 6(4), 273–289. doi: 10.1023/A:1008930821076
- Schooler, J., Gerhard, D. & Loftus, E. (1986). Qualities of unreal. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 171–181. doi: 10.1037/0278-7393.12.2.171
- Sporer, S. L. & Sharman, S. J. (2006). Should I believe this? Reality monitoring of accounts of self-experienced and invented recent and distant autobiographical events. *Applied Cognitive Psychology*, 20, 837–854. doi: 10.1002/acp.1234
- Steller, M. & Köhnken, G. (1989). Criteria-Based Content Analysis. En D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217–245). New York: Springer-Verlag.
- Steyvers, M. (2002). Multidimensional scaling. In *Encyclopedia of Cognitive Science*. (pp. 1–7). London: MacMillan.
- Strömwall, L. A., Bengtsson, L., Leander, L. & Granhag, P. A. (2004). Assessing children's statements: The impact of a repeated Experience on CBCA and RM ratings. *Applied Cognitive Psychology*, 18, 653–668. doi: 10.1002/acp.1021
- Trankell, A. (1972). *Reliability of evidence*. Stockholm: Rotobekman
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–352. doi: 10.1037/0033-295X.84.4.327
- Tversky, A. & Krantz, D. H. (1970). The dimensional representation and the metric structure of similarity data. *Journal of mathematical psychology*, 7, 572–596. doi: 10.1016/0022-2496(70)90041-6
- Undeutsch, U. (1989). The development of statement reality analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 101–121). Netherland: Kluwer Academic Publishers.
- Vrij, A. (2005). Criteria-Based Content Analysis: A Qualitative Review of the First 37 Studies. *Psychology, Public Policy, and Law*, 11(1), 3–41. doi: 10.1037/1076-8971.11.1.3
- Vrij, A., Akehurst, L., Soukara, S. & Bull, R. (2004). Detecting deceit via analysis of verbal and nonverbal behaviour in children and adults. *Human Communication Research*, 30(1), 8–41. doi: 10.1111/j.1468-2958.2004.tb00723.x
- Wicker, A. W. (1975). An application of the multitrait-multimethod logic to the reliability of observational records. *Personality and Social Psychology Bulletin*, 4, 575–579. doi: 10.1177/014616727500100405
- Zahn-Waxler, C., Robinson, J. L. & Emde, R. N. (1992). The development of empathy in twins. *Developmental Psychology*, 28, 1038–1047. doi: 10.1037/0012-1649.28.6.1038

(Article received: 11-03-2014; revised: 13-09-2015; accepted: 24-11-2015)